Zeitschrift: Commentarii Mathematici Helvetici

Herausgeber: Schweizerische Mathematische Gesellschaft

Band: 4 (1932)

Artikel: Sur la définition de la probabilité.

Autor: Dumas, S.

DOI: https://doi.org/10.5169/seals-5615

Sur la définition de la probabilité

par S. DUMAS, Berne

Position de la question

Dans sa définition pascalienne, la probabilité d'un événement est égale au quotient du nombre des cas favorables à l'arrivée de cet événement par le nombre total des cas possibles, en supposant que tous les cas sont également possibles. Tous les auteurs qui utilisent cette définition signalent les critiques qu'elle soulève, mais leurs réponses laissent à désirer; c'est pourquoi nous essayons de fonder le calcul des probabilités en évitant tout cercle vicieux.

Le mathématicien est libre dans ses définitions; nous en donnerons une de la probabilité sans rechercher si d'autres conduiraient à de bons résultats; d'Alembert nous a montré que ce n'est pas toujours vrai; il ne distingue que deux cas, suivant que l'événement se produit ou non. Avec lui, le calcul des probabilités se réduirait à une seule proposition: la probabilité de tout événement est égale à $\frac{1}{2}$. Il n'offrirait aucun intérêt.

Sans ses applications, il est douteux que le calcul des probabilités eût dépassé les mathématiques amusantes. C'est la raison pour laquelle on a recherché dans des considérations physiques une base pour la définition de la probabilité; ces tentatives ont provoqué des travaux de valeur, parmi lesquels nous ne citerons qu'un des derniers, le traité de M. de Mises[1]). Nous craignons que les auteurs de ces ouvrages ne confondent deux questions: la définition mathématique et la définition physique de la probabilité; il faut les élucider l'une et l'autre.

Lorsque nous nous occuperons de la définition physique, nous verrons que ce n'est pas seulement la notion d'égale possibilité qui présente de grosses difficultés, mais déjà celle de cas. Nous montrerons aussi l'importance de la formule de Bayes. Nous verrons que si l'on peut éviter le cercle vicieux dans la définition mathématique, nous ne le pouvons pas dans la définition physique, mais nous le présentons autrement qu'on ne l'a fait jusqu'ici.

---

1) *von Mises:* Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik. Leipzig und Wien. Franz Deuticke 1931.

Nous supposerons premièrement que le nombre des événements est fini; c'est de beaucoup le problème le plus important; puis, nous passerons au nombre illimité d'événements possibles; la question capitale sera de définir le choix au hasard d'un élément.

La définition mathématique de la probabilité

Collection et ordre de multiplicité. Considérons trois éléments a, b, c; nous leur attribuons respectivement les trois entiers positifs l, m et n que nous nommons leurs ordres de multiplicité; nous appelons cet ensemble une collection; nous la désignons par le symbole $\left\{ \begin{matrix} a, & b, & c \\ l, & m, & n \end{matrix} \right\}$; nous appelons C la collection particulière ainsi définie.

Dans la suite, les mots «éléments» et «événements» sont synonymes ainsi que les expressions «ordre de multiplicité» et «nombre de cas favorables à l'arrivée de l'événement».

Pour simplifier l'écriture, nous ne considérons qu'un petit nombre d'éléments; on généralise sans peine nos résultats pour autant que le nombre des éléments et leurs ordres de multiplicité restent finis. Il suffit presque toujours de n'en prendre que deux a et b, autrement dit a et *non-a*.

Nous devons montrer la signification des ordres de multiplicité; à cet effet, à la place de la collection C, envisageons l'ensemble de lettres $(a_1, a_2, a_3 \ldots a_l;\ b_1, b_2, b_3 \ldots b_m;\ c_1, c_2, c_3 \ldots c_n)$; formons tous les arrangements N à N avec répétition de ces $(l + m + n)$ lettres; effaçons les indices; il nous reste des arrangements avec répétition des trois lettres a, b et c; chaque arrangement apparaît plusieurs fois, mais pas en même nombre lorsque l, m et n sont inégaux. Pour rappeler ces différences et pour marquer l'origine de ces arrangements, nous les appelons des «dispositions» N à N issues de la collection C.

Nous considérons comme identiques deux dispositions dans lesquelles les lettres a, b et c se suivent dans le même ordre même si elles proviennent de deux arrangements différents des lettres $a_1, a_2, a_3, \ldots c_n$.

Lorsque nous aurons besoin de distinguer entre les éléments a, b, c et les éléments $a_1, a_2, a_3 \ldots c_n$, nous appellerons les premiers «éléments agrégés» et les seconds «éléments unitaires».

Probabilité. Nous appelons probabilité des éléments a, b et c dans la collection C, les nombres

$$p_a = \frac{l}{l+m+n}; \qquad p_b = \frac{m}{l+m+n}; \qquad p_c = \frac{n}{l+m+n}.$$

Nous pouvons multiplier les ordres de multiplicité par un même nombre ou les diviser par un facteur commun sans modifier les probabilités; les problèmes où cette condition ne pourrait être maintenue n'appartiendraient pas au calcul des probabilités.

Problème fondamental. Ecrivons le tableau complet des $(l+m+n)^N$ dispositions N à N issues de la collection C; définissons deux propriétés relativement à l'ordre dans lequel les lettres a, b et c se succèdent dans une disposition ou relativement à leurs nombres respectifs. Nous désignons par A les dispositions qui ont les deux propriétés et par B celles qui ont la première sans la seconde. Soit L le nombre des dispositions A et M celui des dispositions B. Comme nous l'avons fait pour les éléments a, b et c, nous pouvons distinguer par des indices les diverses dispositions $A_1, A_2, A_3 \ldots A_L$; $B_1, B_2, B_3 \ldots . B_M$. A et B sont les éléments agrégés et $A_1, A_2 \ldots B_M$ les éléments unitaires d'une nouvelle collection $\left\{\begin{matrix} A, & B \\ L, & M \end{matrix}\right\}$. Le problème fondamental consiste à calculer la probabilité $\frac{L}{L+M}$ de l'élément A dans cette collection; on l'appelle la probabilité pour que la disposition ait la seconde propriété lorsqu'on sait qu'elle a la première.

Lorsque dans la collection C, on multiplie les ordres de multiplicité l, m et n par un facteur donné k, le nombre total des dispositions se multiplie par k^N; il en est de même des nombres L et M; la probabilité $\frac{L}{L+M}$ n'est pas modifiée. Un raisonnement semblable amène au même résultat lorsqu'on divise l, m et n par un facteur commun. Le problème fondamental reste dans les limites posées au calcul des probabilités.

Théorème des probabilités totales. Considérons les lettres b et c comme un seul élément d; son ordre de multiplicité est $m+n$ et sa probabilité

$$p_d = p_b + p_c.$$

Théorème des probabilités partielles. Ne considérons plus que les éléments a et b; leurs probabilités sont:

$$p'_a = \frac{l}{l+m} = \frac{p_a}{p_a + p_b}$$

$$p'_b = \frac{m}{l+m} = \frac{p_b}{p_a + p_b} \, .$$

Ce sont les probabilités de a et b lorsqu'on sait que c ne s'est pas présenté.

Théorème des probabilités composées. Dans une collection de deux éléments a et b, la probabilité de a est égale à p_a ; nous combinons a et b avec deux autres éléments c et d et nous désignons les éléments composés par (ac), (ad), (bc) et (bd); nous formons avec ces quatre éléments une collection qui doit satisfaire à deux conditions:

1° la probabilité pour qu'un élément de la nouvelle collection commence par a est égale à p_a, probabilité de a dans la première collection;

2° la probabilité pour que le second élément d'un élément composé soit c lorsqu'on sait que le premier est a est égale à un nombre donné $p_{c|a}$.

On demande quelle est la probabilité p_{ac} de l'élément (ac) dans la nouvelle collection.

Désignons par p_{ad} la probabilité de l'élément (ad) dans la nouvelle collection.

En vertu du théorème des probabilités totales, la première condition nous donne

$$p_{ac} + p_{ad} = p_a \, .$$

La seconde condition, en vertu du théorème des probabilités partielles, peut s'écrire:

$$p_{c|a} = \frac{p_{ac}}{p_{ac} + p_{ad}} \, .$$

Si nous éliminons p_{ad} entre ces deux égalités, nous obtenons:

$$p_{ac} = p_a \cdot p_{c|a} \, .$$

Afin de définir complètement la nouvelle collection, il faudrait encore connaître la probabilité $p_{c|b}$ pour que le second élément d'un élément composé soit c lorsqu'on sait que le premier est b; il y a beaucoup de questions qu'on peut résoudre sans cela.

Lorsque

$$p_{c|a} = p_{c|b}$$

on dit que l'événement c est indépendant de l'arrivée des événements a ou b; dans le cas contraire, on dit qu'il en dépend.

Théorème de Bernoulli. La définition de l'espérance mathématique, ou valeur probable d'une fonction, n'offre pas de difficultés; nous disposons donc des bases nécessaires pour démontrer les autres théorèmes du calcul des probabilités tels que ceux de Laplace et de Bernoulli, ainsi que la loi des grands nombres. Ils deviennent des théorèmes d'analyse combinatoire; en particulier celui de Bernoulli peut s'énoncer comme il suit:

Soient les éléments a et b dont les ordres de multiplicité sont l et m; nous en formons les $(l+m)^N$ dispositions N à N; désignons par s le nombre des lettres a contenues dans une disposition et par $f(N, \varepsilon)$ le nombre des dispositions pour lesquelles

$$-\varepsilon < \frac{l}{l+m} - \frac{s}{N} < \varepsilon;$$

quel que petit que soit le nombre positif ε, nous avons

$$\lim_{N=\infty} \frac{f(N, \varepsilon)}{(l+m)^N} = 1.$$

Les probabilités inverses. Considérons k collections de la forme $\begin{Bmatrix} a, & b \\ l_i, & m_i \end{Bmatrix}$, i prenant les valeurs 1, 2, 3 ... k et désignons-les par C', C'', C''', ... $C^{(k)}$. Nous les prenons comme éléments d'une nouvelle collection C dans laquelle $C^{(i)}$ a l'ordre de multiplicité de M_i. Nous devons définir le tableau complet issu de la collection C des dispositions N à N des lettres a et b.

Le tableau complet des dispositions N à N issues de la collection $C^{(i)}$ contient $(l_i + m_i)^N$ éléments; ce nombre est fonction de i; nous pouvons éviter cette dépendance en multipliant les ordres de multiplicité l_i et m_i par $\frac{1}{l_i + m_i} \prod_{i=1}^{k} (l_i + m_i)$. Le nouveau tableau issu de la collection $C^{(i)}$ contient $\prod_{i=1}^{k} (l_i + m_i)^N$ dispositions.

Pour tenir compte de l'ordre de multiplicité M_i, nous écrivons M_i fois le nouveau tableau ; nous procédons ainsi pour toutes les valeurs de i et nous obtenons $(M_1 + M_2 + M_3 + \ldots + M_k)$ tableaux de $\prod_{i=1}^{k} (l_i + m_i)^N$ dispositions. Leur ensemble constitue, par définition, le tableau complet issu de la collection C des dispositions N à N des lettres a et b.

Considérons une de ces dispositions et désignons-la par A. Supposons qu'elle figure S_i fois dans l'ensemble des M_i tableaux issus de la collection $C^{(i)}$; elle apparaît $\sum_{i=1}^{k} S_i$ fois dans le tableau complet issu de la collection C. On appelle

$$P_i = \frac{S_i}{\sum_{i=1}^{k} S_i}$$

la probabilité pour que la disposition A soit issue de la collection $C^{(i)}$; si nous employions le langage usuel, nous dirions que c'est la probabilité pour que la disposition A soit due à la cause $C^{(i)}$.

Il est utile de mettre P_i sous une autre forme.

La probabilité de $C^{(i)}$ dans la collection C est égale à

$$\pi_i = \frac{M_i}{\sum_{i=1}^{k} M_i};$$

c'est ce qu'on appelle généralement la probabilité à priori de la cause $C^{(i)}$.

Désignons par p_i la probabilité de la disposition A dans le tableau issu de la collection $C^{(i)}$; c'est la probabilité pour que la cause $C^{(i)}$, lorsqu'elle agit, provoque l'événement A. Nous avons :

$$p_i = \frac{S_i}{M_i \prod_{i=1}^{k} (l_i + m_i)^N}.$$

Une élimination entre les trois dernières égalités nous donne

$$P_i = \frac{p_i \pi_i}{\sum_{i=1}^{k} p_i \pi_i}.$$

C'est la formule de Bayes.

Schéma des urnes. Pour soutenir la pensée, il est commode de recourir au schéma des urnes. Nous supposons une urne contenant l boules blanches, m noires et n rouges. Nous les tirons, en supposant que l'on remette toujours la boule extraite dans l'urne et qu'on mélange chaque fois le tout; si le nombre des tirages est N et si nous notons respectivement par a, b et c la sortie des boules blanches, noires et rouges, nous écrivons une disposition N à N de la nature de celles qui sont issues de la première collection que nous avons envisagée. Nous pouvons donner ainsi une représentation des questions ressortissant au calcul des probabilités. Comme une figure géométrique, une urne réelle n'est jamais parfaite; nous supposons la nôtre idéale, ce qui signifie que si nous procédons un très grand nombre de fois à un tirage de N boules, toutes les dispositions apparaîtront en nombres proportionnels à leur ordre de multiplicité dans le tableau complet des dispositions N à N issues de la collection $\left\{\begin{matrix} a, & b, & c \\ l, & m, & n \end{matrix}\right\}$.

Nous disons que la boule extraite a été «désignée par le hasard». Nous devons avoir un schéma physique pour donner un sens à cette expression. Au point de vue mathématique, les éléments de nos collections sont équivalents et nous ne pouvons pas en choisir l'un plutôt que l'autre. Toutefois, comme une réalisation physique n'est jamais qu'approximative, le mathématicien ne dispose d'aucun moyen parfait de prendre un objet au hasard.

Dans les probabilités inverses, la représentation est la suivante. Nous supposons $(M_1 + M_2 + M_3 + \dots + M_k)$ urnes; M_i d'entre elles portent un numéro i invisible pour l'expérimentateur; elles contiennent l_i boules blanches et m_i noires; l'expérimentateur prend une des urnes au hasard; il en extrait, en remettant toujours la boule sortie, s boules blanches et t noires; étant donnés les nombres s et t, on demande quelle est la probabilité pour que l'urne d'où sortent les boules porte le numéro i.

Le schéma des urnes est très souple; quelquefois, il vaut pourtant mieux recourir à une autre représentation: jeu de pile ou face, dés, cartes, etc. Le fonds de la question n'est pas modifié.

Un problème. Une urne contient des boules blanches et noires dans une proportion inconnue; on fait $(s + t)$ tirages, en remettant toujours la boule extraite dans l'urne; il sort s boules blanches et t noires. Nous faisons une nouvelle suite de $(u+v)$ tirages; quelle est la probabilité pour que u des boules extraites soient blanches et v noires?

Ce problème est insoluble, faute de données nécessaires. On peut cependant faire les considérations suivantes:

Nous supposons que l'urne considérée appartient à l'ensemble des urnes que nous venons d'envisager; sa composition est telle que la boule blanche a une probabilité de sortie égale à h_i; la probabilité à priori de cette composition est égale à π_i. Nous pouvons appliquer la formule de Bayes avec

$$p_i = \frac{(s+t)!}{s!\ t!} h_i^s (1-h_i)^t$$

et nous obtenons pour la probabilité à posteriori de la dite composition

$$P_i = \frac{\pi_i h_i^s (1-h_i)^t}{\sum_{i=1}^{k} \pi_i h_i^s (1-h_i)^t}.$$

Considérons le produit $h_i^s(1-h_i)^t$ dans l'intervalle $(0, 1)$ et supposons s et t très grands. Lorsque h_i est égal à 0, le produit s'annule s fois; il croît très lentement, puis extrêmement rapidement, passe par un maximum pour

$$h_i = \frac{s}{s+t} = h,$$

décroît extrêmement rapidement, puis très lentement et s'annule t fois lorsque h_i est égal à 1. Supposons s et t assez grands pour que par rapport au plus grand terme $h^s(1-h)^t$ on puisse négliger tous les autres. Les deux termes de P_i sont égaux; nous avons la certitude que la probabilité de sortie de la boule blanche est égale à h. Autrement dit, on prend pour probabilité de sortie de la boule blanche la valeur la plus probable et l'on néglige les autres.

Ce raisonnement suppose que la valeur de π_i qui correspond à h soit différente de zéro; autrement dit, si B est le nombre des boules contenues dans l'urne, $\frac{s}{s+t} B$ doit être un nombre entier. En outre, il faut qu'aucun des π_i ne soit assez grand pour nous obliger à conserver les termes que nous avons négligés.

Ces hypothèses sont très osées. Nous ignorons le nombre B des boules contenues dans l'urne. Nous ignorons aussi comment elle a été composée, ce qui fait que nous n'avons aucune indication sur la valeur des π_i; nous sommes en face d'une seule urne; nous ne pouvons donner à π_i aucune signification.

Si nous passons sur ces objections, la probabilité pour qu'une nouvelle série d'épreuves nous amène u boules blanches et v noires est

$$P_{u,v} = \frac{(u+v)!}{u!\ v!} h^u (1-h)^v = \frac{(u+v)!}{u!\ v!} \frac{s^u\ t^v}{(s+t)^{u+v}}.$$

Lorsqu'on veut conserver les probabilités autres que h, on raisonne comme il suit: en vertu des théorèmes des probabilités totales et composées, la probabilité d'extraire u boules blanches et v noires est égale à

$$\begin{aligned} P'_{u,v} &= \sum_{i=1}^{k} P_i \frac{(u+v)!}{u!\ v!} h_i^u (1-h_i)^v \\ &= \frac{(u+v)!}{u!\ v!} \frac{\sum_{i=1}^{k} \pi_i h_i^{s+u} (1-h_i)^{t+v}}{\sum_{i=1}^{k} \pi_i h_i^s (1-h_i)^t}. \end{aligned}$$

On admet que les π_i sont tous égaux entre eux et l'on évalue les deux termes de la fraction au moyen des intégrales eulériennes de première espèce; on trouve

$$P'_{u,v} = \frac{(u+v)!\ (s+u)!\ (t+v)!\ (s+t+1)!}{u!\ v!\ (s+t+u+v+1)!\ s!\ t!}$$

en particulier, si nous posons

$$\begin{cases} u = 1 \\ v = 0 \end{cases}$$

c'est-à-dire, si nous cherchons la probabilité pour que la prochaine boule extraite de l'urne soit blanche, nous trouvons:

$$P'_{1,0} = \frac{s+1}{s+t+2}\ ^{2)}.$$

---

[2]) Nous avons traité ce problème plus en détail dans: Les probabilités inverses et la construction des tables de mortalité. Festgabe Moser. Bern. Kommissionsverlag Stämpfli und Cie 1931.

Les deux raisonnements présentent des difficultés équivalentes; dans les deux cas, nous faisons des hypothèses sur les π_i que nous ne connaissons pas. Si dans le premier, nous admettons que $\frac{s}{s+t} B$ est un nombre entier, dans le second, nous devrions savoir combien l'urne contient de boules pour évaluer l'erreur commise en remplaçant des sommes finies par des intégrales. Si s et t sont grands par rapport à B, il faut craindre que l'approximation ne soit mauvaise.

La définition physique de la probabilité

Cette définition se heurte à des difficultés insurmontables en logique. Nous ne les croyons pourtant pas essentiellement différentes de celles qu'on rencontre chaque fois qu'on applique les mathématiques à la physique. La logique ne nous permet pas de conclure; nous prenons confiance lorsque l'expérience confirme nos prévisions. Au fond, la notion de cause n'est pas plus facile à comprendre que celle de hasard; nous y sommes plus habitués.

Le hasard. On ne peut appliquer le calcul des probabilités qu'aux phénomènes dus au hasard.

Nous ne rappellerons pas ici, faute de place, le résultat des nombreuses études consacrées à la nature du hasard; elles nous permettent, quelquefois même avant d'avoir observé un phénomène, de présumer s'il est fortuit ou dû à des causes déterminées; nous vérifions expérimentalement cette présomption. Le moyen en est de confronter l'expérience avec les prévisions du calcul des probabilités. Nous tournons dans un cercle vicieux: le calcul des probabilités s'applique aux phénomènes dus au hasard; on reconnaît le hasard au fait qu'on peut appliquer le calcul des probabilités aux phénomènes qu'il provoque.

Si nous connaissons le résultat d'une expérience sans savoir comment on l'a faite, nous ne pouvons pas, en toute rigueur, exclure le hasard ou, au contraire, affirmer qu'il est intervenu. Nous trouvons un jeu de cartes dans un ordre fixé d'avance; il est possible, quoique invraisemblable, qu'une personne ait battu le jeu et soit arrivée au dit ordre sans le vouloir. Considérons aussi une suite de N nombres entiers, au moins égaux à 0, au plus égaux à 36; désignons le $i^{\text{ième}}$ par n_i; nous pouvons toujours définir une fonction $f(i)$ telle que

$$f(i) = n_i;$$

notre suite peut être une permanence à la roulette ou la table numérique de la fonction $f(i)$.

M. de Mises[3]), considérant des suites de nombres, pose comme critère du hasard, leur désordre complet; cette notion est intéressante; elle ne peut cependant pas nous satisfaire. Si la suite est finie, on peut toujours imaginer une loi de succession; d'autre part, une suite infinie ne peut résulter d'aucune expérience; pour la donner, il faut la définir, c'est-à-dire donner la loi des termes; de plus, l'étude de cette loi nous fait quitter le domaine de la physique pour rentrer dans celui des mathématiques.

Le hasard ou la causalité dépendent dans une large mesure de la question, de l'échelle ou du point de vue.

Dans leur suite naturelle, les nombres impairs et pairs se succèdent régulièrement; si nous extrayons une boule d'une urne qui en contient n numérotées de 1 à n, c'est par hasard que le numéro sorti est pair ou impair.

Nous nous efforçons de connaître les modifications de l'atome par le principe de causalité et nous en déduisons des résultats d'ensemble par le calcul des probabilités.

On admet que les atomes sont des systèmes solaires en miniature; s'ils sont habités par des êtres qui ont la même logique que nous, leurs savants traitent du mouvement moléculaire par nos méthodes astronomiques. Symétriquement, on peut concevoir le ciel étoilé comme une bulle de gaz dans le laboratoire d'un physicien énorme qui en fait la théorie cinétique.

L'accident illustre bien l'importance du point de vue. Les intéressés en étudient toutes les causes pour fixer les responsabilités, tandis que le statisticien le considère comme un événement fortuit. Cette double attitude est particulièrement nette chez le directeur d'une compagnie d'assurances; d'abord, il examine les circonstances de l'accident pour voir s'il doit indemniser la victime, puis il fait entrer le sinistre dans ses statistiques pour comparer ses primes aux indemnités et pour évaluer ses risques futurs.

Premier exemple. Nous supposons, comme dans le problème ci-dessus, que nous avons extrait d'une urne s boules blanches et t noires; seulement, il s'agit d'une urne non pas idéale mais réelle. Nous disons que la probabilité d'extraire une boule blanche de l'urne est égale à $\frac{s}{s+t}$.

[3]) Wahrscheinlichkeit, Statistik und Wahrheit. Wien. Julius Springer 1928.

Nous assimilons l'urne physique à l'urne mathématique; nous appliquons la formule de Bayes; à cet effet, nous admettons que nous pouvons éliminer ou négliger les probabilités à priori π_i; cette supposition est nécessaire pour le calcul et pour éviter qu'il faille connaître des probabilités pour définir la probabilité. En outre, comme nous le verrons dans notre second exemple, il est souvent impossible de donner aux π_i une signification physique.

Nous posons la probabilité d'extraire une boule blanche égale à cette probabilité lorsque l'urne a la composition la plus probable; nous négligeons ainsi complètement les autres compositions.

Nous avons raisonné jusqu'ici sans connaître le nombre des boules contenues dans l'urne; il arrive que nous sachions qu'elle en renferme S blanches et T noires. La question n'est pas essentiellement autre. Distinguons deux cas:

Premièrement, la probabilité observée $\frac{s}{s+t}$ d'extraire une boule blanche est voisine de la probabilité présumée $\frac{S}{S+T}$; nous admettrons plus facilement que le hasard intervient et que le calcul des probabilités est applicable, que lorsque nous ignorons tout sur la composition de l'urne.

Secondement, les probabilités présumée et observée sont très différentes l'une de l'autre. Nous aurons de la peine à conclure; nous étudierons la possibilité d'une cause jusqu'alors inconnue. Nous pouvons aussi nous trouver en présence d'une disposition très rare; c'est une des difficultés que l'on rencontre dans les applications du calcul des probabilités. Nous tombons sur un cas exceptionnel qui met nos prévisions en défaut; il n'infirme pas la théorie car les cas les plus rares doivent se présenter quelquefois.

Nous schématisons en n'envisageant que deux cas; en réalité, il y a toutes les possibilités intermédiaires. Suivant les circonstances, notre raisonnement sera plus ou moins satisfaisant; il ne sera jamais strictement rigoureux. Si dans une longue suite de tirages, les probabilités observée et présumée restaient constamment très voisines l'une de l'autre, nous nous demanderions si une influence régularisatrice ne diminue ou même ne supprime pas le hasard.

Remarquons encore que si nous ignorons le nombre de boules contenues dans l'urne, nous ne connaissons pas le nombre des cas possibles. La difficulté que présente la définition classique ne consiste pas seulement à savoir si tous les cas sont également possibles; elle est plus profonde car nous ne pouvons pas énumérer les cas.

Second exemple. Nous reprenons ici un exemple que nous avons traité plus en détail dans notre mémoire précité; celui de la détermination des probabilités annuelles de décès. Si nous n'en conservons que les éléments essentiels, la méthode en est la suivante.

On observe pendant une année un groupe de personnes d'âge x; si leur nombre est λ_x et qu'il se produise δ_x décès dans le cours de l'année, on dit que la probabilité annuelle de décès pour des personnes d'âge x est

$$q_x = \frac{\delta_x}{\lambda_x}.$$

Ramenons le procédé au schéma des urnes. D'une urne dont la composition est inconnue, nous tirons λ_x boules en remettant toujours la boule extraite dans l'urne; δ_x boules extraites sont blanches, les autres noires. Nous en concluons que la probabilité de sortie d'une boule blanche est égale à $\frac{\delta_x}{\lambda_x}$.

Nous sommes en présence d'une seule urne, sur la composition de laquelle nous n'avons aucun renseignement; nous ignorons complètement la valeur des probabilités à priori π_i et nous n'en pouvons donner aucune interprétation physique.

Nous ne pouvons pas davantage fixer le nombre des cas possibles. Si pour un homme de trente ans, la probabilité annuelle de mourir dans l'année est égale à 0,006, cela ne peut signifier qu'il faut se représenter pour cet individu 1000 cas différents et qu'il mourra dans 6 d'entre eux.

On confond généralement le nombre des cas possibles avec celui des cas observés, autrement dit, le nombre des boules contenues dans l'urne avec celui des boules extraites. Or, nous sommes dans un problème de probabilités inverses; nous connaissons le nombre des cas observés; sur 1000 hommes, 6 sont morts et 994 ont survécu; sur 1000 boules extraites, 6 sont blanches et 994 noires. Pour conclure à la probabilité, il faut passer par la formule de Bayes.

Cet exemple met en évidence une nouvelle condition. Les λ_x personnes doivent constituer un groupe homogène au point de vue de la mortalité; s'il en était autrement, nous devrions envisager plusieurs urnes de compositions différentes; on en extrairait des boules mais on ne connaîtrait que le résultat global de l'expérience. Toute conclusion serait aventurée. Or, dans la réalité, le groupe n'est jamais tout à fait homogène.

Si nous utilisons la probabilité de décès pour évaluer un nombre de morts, le groupe auquel nous l'appliquons doit être homogène à celui que nous avons observé; nous ne devons pas conclure d'une urne à une autre. Cette condition n'est jamais remplie dans la réalité; la mortalité varie avec l'époque.

Nos deux exemples présentent une différence. Dans le premier, nous pouvons noter l'ordre dans lequel les boules blanches et noires se succèdent; dans le second, nous ne connaissons que le résultat global. Cette différence n'influence pas sur la probabilité; en revanche, si dans le premier cas, la suite des couleurs est très irrégulière, nous conclurons plus facilement au hasard; si elle est très régulière, nous penserons, au contraire, à une cause déterminée.

On pourrait multiplier ces exemples; on y trouverait toujours, sous une forme ou sous une autre, des difficultés semblables à celles que nous avons relevées.

La probabilité. Si un événement est dû au hasard et que sur $(s+t)$ épreuves il se soit produit exactement s fois, sa probabilité est égale au quotient $\frac{s}{s+t}$.

Cette définition a des défauts que nous croyons essentiels.

Il faut que le phénomène soit dû au hasard; or nous avons vu que nous ne le savons jamais avec certitude.

Nous justifions notre définition par la formule de Bayes; nous devons donc connaître la probabilité à priori π_i pour que le phénomène ait une probabilité p_i; d'une part, nous nous trouvons dans un cercle vicieux; d'autre part, souvent nous ne pouvons pas donner d'interprétation physique à π_i. Lorsqu'on se base sur l'énumération des cas possibles, on exige qu'ils soient également possibles, autrement dit également probables. Le cercle vicieux est équivalent à celui que nous venons de signaler.

En donnant cette définition, nous postulons que si nous observons plusieurs événements différents, nous pouvons combiner leurs probabilités physiques selon les règles du calcul des probabilités. En particulier, nous admettons que si sur $(u+v)$ nouvelles épreuves, l'événement se produit u fois, nous devons avoir, lorsque u et v croissent indéfiniment

$$\operatorname{Lim} \frac{u}{u+v} = \frac{s}{s+t}.$$

Nous ignorons le nombre $(s + t)$ d'épreuves qu'il faut faire pour pouvoir conclure; il dépend de l'événement considéré; en particulier, il est très grand lorsque la probabilité est très petite.

Puis viennent des difficultés de réalisation; nous devons faire toutes les épreuves dans des conditions identiques et n'appliquer la probabilité obtenue à une nouvelle série d'épreuves que lorsqu'elles sont faites dans les mêmes conditions que les premières.

En logique, pour des raisons inhérentes à l'esprit humain, notre position est faible. Mais, ici comme ailleurs, nous prouvons le mouvement en marchant. Les beaux résultats du calcul des probabilités témoignent de sa valeur; il est un des instruments indispensables de la science moderne. Cependant, nous devons être prudents dans nos applications. Comme toute grandeur physique, la probabilité n'est connue, nous aimerions dire mesurable, qu'avec une certaine approximation; il ne faut pas vouloir en tirer trop. C'est ce que nous avons illustré par un exemple dans le mémoire que nous avons déjà cité. Le calcul des probabilités rend d'éminents services aux actuaires; il leur permet de calculer les primes d'assurances sur la vie et les réserves mathématiques; en revanche, on est allé trop loin en calculant l'erreur probable qui affecte les probabilités annuelles de décès.

Les difficultés logiques auxquelles se heurte l'application sont peut-être un peu plus nombreuses pour le calcul des probabilités et la mécanique statistique que pour l'analyse mathématique et la mécanique rationnelle; elles ne sont pas essentiellement différentes. Dans les deux cas, nous observons des phénomènes; nous en tirons des conclusions et, à l'aide des mathématiques, nous prévoyons d'autres phénomènes. Il arrive toujours un moment où nos prévisions sont démenties par l'expérience; nous devons alors modifier nos prémisses; nous poursuivons jusqu'à ce qu'un nouvel échec nous contraigne à un nouveau changement.

Ignorance. On a beaucoup trop dit que le calcul des probabilités nous permettait d'autant mieux de conclure que notre ignorance était plus grande. Au contraire, la connaissance du phénomène nous est fort utile. C'est parce que nous savons que les dés sont bien faits, que nous pouvons admettre à priori que la probabilité d'amener l'as est voisine de $\frac{1}{6}$ et nous contenter d'une expérience assez courte pour nous convaincre que le dé n'est pas pipé. En revanche, s'il s'agissait d'un polyèdre irrégulier à 6 faces, il nous faudrait un grand nombre de jets pour déterminer la probabilité avec laquelle chaque face apparaît. Notre analyse confirme ce fait. C'est parce que nous ne connaissons pas les probabilités à priori π_i qu'il nous est si difficile de conclure.

Le nombre des événements est illimité

Choix au hasard. Un problème important se pose; nous avons une infinité d'éléments, comment en choisir un au hasard?

Deux exemples se présentent à nous; les éléments peuvent être des points ou des nombres. Aucun moyen physique ne nous permet de désigner un point; nous fixons toujours un petit espace dont la grandeur dépend de nos moyens d'observation. Dans toute région finie, le nombre de ces espaces est fini car, pour en distinguer deux, il faut un déplacement minimum de l'un par rapport à l'autre. En revanche, nous pouvons indiquer un nombre en toute rigueur. Notre problème devient donc: choisir un nombre au hasard dans un ensemble infini.

Les événements sont dénombrables. Soit une infinité d'éléments a_1, a_2, a_3 ... a_i ... ; nous en considérons n et nous formons la collection $\begin{Bmatrix} a_1, & a_2, & a_3 & \dots & a_i & \dots & a_n \\ l_{1,n}, & l_{2,n}, & l_{3,n} & \dots & l_{i,n} & \dots & l_{n,n} \end{Bmatrix}$; les ordres de multiplicité $l_{i,n}$ sont des fonctions de i et de n. Nous faisons croître n indéfiniment; si elle existe, la limite

$$\lim_{n=\infty} \frac{l_{i,n}}{\Sigma\, l_{i,n}} = p_i$$

est la probabilité de a_i dans l'ensemble infini considéré.

Il peut arriver que tous les $l_{i,n}$ restent finis; tous les p_i sont égaux à zéro. Dans ce cas, nous n'avons pas de procédé physique pour désigner un élément au hasard. Si tous les p_i ne sont pas nuls, nous pouvons procéder comme il suit:

Soit une suite de nombres rationnels et positifs p_1, p_2, p_3 ... p_i ... tels que

$$p_1 + p_2 + p_3 + \dots + p_i + \dots\dots = 1.$$

Nous constituons une suite d'urnes numérotées 1, 2, 3 ... i ... ; nous mettons dans chacune d'elles des boules blanches et noires dans une proportion telle que la sortie d'une boule blanche de l'urne i ait une probabilité

$$P_i = \frac{p_i}{1 - p_1 - p_2 - p_3 - \dots - p_{i-1}}.$$

Nous le pouvons car de p_i positif et de

$$1 - p_1 - p_2 - p_3 - \dots - p_i > 0$$

résulte

$$0 < P_i < 1.$$

Nous extrayons une boule de la première urne; si elle est blanche, c'est le nombre 1 qui est désigné au hasard. Si elle est noire, nous extrayons une boule de la seconde urne; si elle est blanche, c'est le nombre 2 qui a été désigné. Si elle est noire, nous tirons une boule de la troisième urne; nous poursuivons ainsi les extractions jusqu'à ce qu'il sorte une boule blanche; si elle sort de l'urne i, c'est le numéro i qui a été désigné par le hasard. Si les urnes sont composées comme nous l'avons dit, la probabilité de désigner le nombre i est égale à p_i.

C'est exact pour i égal à 1; par définition la probabilité d'extraire une boule blanche de l'urne 1 est égale à p_1. C'est encore exact pour i égal à 2; il faut que le premier tirage ait amené une boule noire, ce qui a la probabilité $(1 - p_1)$ et le second une boule blanche, ce qui a la probabilité $\frac{p_2}{1 - p_1}$; le produit est égal à p_2.

Si notre proposition est vraie pour les valeurs 1, 2, 3 ... $i - 1$, elle est encore vraie pour i. Pour que le nombre i soit désigné, il faut qu'aucun des nombres plus petits ne l'ait été, ce qui a la probabilité $1 - p_1 - p_2 - p_3 - \dots - p_{i-1}$; il faut de plus qu'une boule blanche sorte de l'urne i ce qui a pour probabilité P_i; le produit est égal à p_i, ce qui démontre notre affirmation.

Les événements ne sont pas dénombrables. Soit $f(x)$ une fonction intégrable dans l'intervalle (a, b); on appelle probabilité pour qu'un nombre de cet intervalle soit compris entre x_0 et x, l'expression

$$p(x_0, x) = \frac{\int_{x_0}^{x} f(x)\,dx}{\int_a^b f(x)\,dx}.$$

La probabilité dépend ainsi du choix de la fonction $f(x)$.

Cette définition n'offre aucune difficulté mathématique. On la généralise aisément, si l'élément considéré, au lieu d'être un nombre unique, consiste en plusieurs nombres comme les coordonnées d'un point situé dans un espace à plusieurs dimensions; on arrive à une intégrale multiple. La notion d'intégrale est aussi prise dans toute sa généralité; c'est ainsi qu'on utilise souvent les intégrales de Stieljes.

En revanche, il n'existe aucun procédé de choisir au hasard un élément dans une infinité non dénombrable. Cette remarque est importante pour les applications. Les intégrales qu'on y rencontre doivent être considérées comme les expressions approchées de sommations finies. Le passage à la limite n'offre guère de difficultés; il faut toutefois être attentif à la manière de poser le problème dans le domaine discontinu; sinon, on risque de tomber dans les erreurs contre lesquelles le paradoxe de Bertrand doit nous mettre en garde.

(Reçu le 30 janvier 1932)